

Table A.1: Ranking of segmentation results across the datasets

Model	Dataset	Liv	Spl	LKid	RKid	Aor	Sto	IVC	Pan	Eso	Gall	RAG	LAG
nnUNet	BTCV	1	2	3	6	5	4	7	8	9	12	10	11
	AMOS	1	2	4	3	5	6	7	9	10	8	12	11
	TotalSeg	1	2	5	4	3	6	7	9	8	12	10	11
	Average Rank	1.00	2.00	4.00	4.33	4.33	5.33	7.00	8.67	9.00	10.67	10.67	11.00
Swin-UNETR	BTCV	1	3	6	5	2	4	7	8	9	10.5	10.5	12
	AMOS	1	2	4	3	5	6	7	9	10	8	11	12
	TotalSeg	1	2	5	4	3	6	7	8	9	12	11	10
	Average Rank	1.00	2.33	5.00	4.00	3.33	5.33	7.00	8.33	9.33	10.17	10.83	11.33